DOCUMENT RESUME

ED 254 536 TM 850 042

AUTHOR Schoen, Frederick, E.; Childs, John W.
TITLE Formative Evaluation of Instructional Products: Is It Worth It?
PUB DATE 81
NOTE 17p.
PUB TYPE Viewpoints (120) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Elementary Secondary Education; *Evaluation Criteria; Evaluation Methods; *Formative Evaluation; Instructional Development; *Instructional Materials; *Material Development; Models; Program Evaluation; *Validity
IDENTIFIERS Evaluation Problems; Scriven (RC)

ABSTRACT

Formative evaluation of instructional materials may not be time efficient or cost effective. Scriven's definition of formative evaluation, used in this paper, applies to the assessment and refinement of instructional materials still under development. Implementation of formative evaluation techniques does not seem to yield significant improvement in learner performance after the first draft stage. Formative evaluation is difficult to validate and lacks generalizability. A simulation game, Return on Investment (ROI), was developed to assess the effectiveness of formative evaluation procedures. The formative evaluation model developed to revise the game was determined to be inadequate because no significant difference in learner performance was produced by the two versions of ROI. This was a further indication that the utility of formative evaluation as an instructional development technique is questionable. (DWH)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED254536

# FORMATIVE EVALUATION OF INSTRUCTIONAL PRODUCTS

## Is it worth it?

Submitted by:

Frederick E. Schoen, Ph.D.

Instructor, Wayne State University
Teacher, Detroit Public Schools
32341 Dover
Warren, Michigan 48093

Home: (313) 268-0292
Ofc.: (313) 577-1740

and

John W. Childs, Ph.D.

Professor
Instructional Technology
53 College of Education
Wayne State University
Detroit, Michigan 48202

Home: (313) 557-8915
Ofc.: (313) 577-1740

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

FE Schoen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 850 092

Please direct all correspondence regarding this manuscript to Dr. Frederick E. Schoen at address shown above.

# FORMATIVE EVALUATION OF INSTRUCTIONAL PRODUCTS

## Is it worth it?

The idea of improving instructional products through formative evaluation possesses a certain seductive logic that is difficult to refute. Most instructional developers believe that an interactive evaluation process ought to result in superior instructional products. It is the purpose of this article to suggest that uncritical acceptance of the efficacy of formative evaluation may lure the unwary instructional developer into a perilous strait between the Scylla of time-inefficiency and the Charybdis of cost ineffectiveness. For the instructional designer/developer with a modicum of experience and sophistication, formative evaluation of instructional products may not be worth the time and effort.

Before advancing evidence to support this contention, it is appropriate to explain what is meant by formative evaluation, particularly in the context of the present discussion. Although the concept of formative evaluation can be traced at least as far back as the 1920's (Cambre, 1981), its widespread application to the development of instructional systems is relatively recent. Scriven (1967) first coined the phrase, applying it to the assessment and refinement of instructional materials still under development. He operationalized it by recommending that curriculum de-

signers subject their products to tryouts with selected samples of learners and teachers.

Scriven went beyond learner testing in the formative evaluation process. He also regarded comments by teachers, criticism by subject specialists, and interest and attitudinal reactions by students as relevant sources of information for formative evaluation (Bloom, Hastings, & Madaus, 1971).[1] Subsequently, other theorists have supplied their own definitions of formative evaluation without materially altering Scriven's original concept (Abedor, 1972; Baker, 1974; Cunningham, 1972; Dick, 1977; Light & Reynolds, 1972; Sanders, 1972; Stake, 1969, 1972).

Similarly, a number of theoretical models have been advanced in an attempt to put some practical meat on the conceptual bones of formative evaluation (Abedor, 1971; Baker & Alkin, 1973; Etlinger, 1975; Nevo, 1976; Sanders & Cunningham, 1972, 1973; Tennyson, 1976; Thiagarajan, 1976). But application of any of the existing formative evaluation

---

[1]Bloom (1968) borrowed the phrase formative evaluation and incorporated it into his concept of mastery learning. In his frame of reference, formative evaluation connoted tests and measures used in a course to improve student learning, i.e., a technique for diagnosing student progress (Miree & Kibler, 1977).

models draws attention to their inadequacies as tools of the professional instructional developer.

A serious problem exists; there is little hard data to indicate that implementation of any of the schemes of formative evaluation yield significant improvement in learner performance after the first draft stage. Since too many first drafts become final drafts, there is some merit in even the most rudimentary evaluation procedure. But almost any kind of systematic evaluation procedure on first draft materials (even one based on the developer's intuition) should generate significant performance improvement. Claims for formative evaluation based on first draft data represent deck-stacking, pure and simple. The findings are about as surprising as discovering that Jack Nicklaus can out-drive the average weekend golfer.

All this is to say that formative evaluation suffers from a serious validation problem. Claims for various formative evaluation procedures have far outdistanced the evidence available to support the effectiveness of those procedures. Smith and Murray's (1975) assertion still holds:

> In evaluation . . . there is not only a lack of research evidence on the utility of various models or conceptual approaches, but a lack of evidence on the effectiveness of component procedures and the basic assumptions within the models themselves. It appears

> . . . that these models cannot be considered anything more than convenient conceptual mechanisms or, at most, value based (not empirically based) claims about what constitutes "good" . . . evaluation procedures. (p. 15)

Questions of validity aside, a second serious weakness of the various approaches to formative evaluation is their lack of generalizability. Gooler's contention that "formative evaluation is not meant to be generalizable" (1980, p. 7) notwithstanding, a "model" for evaluation which is applicable only to a single instructional product is really no model at all and contributes little to theory building. It seems reasonable to expect that a formative evaluation process, to be considered effective, ought to satisfy a second criterion: in addition to resulting in an instructional product which brings about demonstrable improvement in learner performance, it also ought to be applicable to a broad range of products which are in various stages of development.

To test this second proposition, a modest experiment was conducted (Schoen, 1980). Beginning with a role-play simulation game entitled Return on Investment (ROI) (Schoen, 1977), a formative evaluation procedure was devised which was used to create a second version of the game. The success of the procedure was to be measured by the degree to which the revised game produced improvement in learner

achievement of ROI's cognitive and affective objectives. If the formative evaluation model proved valid (i.e., produced a second version of ROI which brought about significant improvement in learner performance), it would then "qualify" to be tested on other simulation games to determine its generalizability.

The simulation game chosen for evaluation, Return on Investment, was intended to increase player knowledge of, and empathy for, the complexities involved in business decision-making. It was selected for two reasons. First, it had been created intuitively without regard to any thought of formative evaluation; and second, it had been used with classroom groups on numerous occasions, had proven playable and, as a result, was said to be in an advanced stage of development. The formative evaluation procedure which evolved during several preliminary playings of ROI was based on the Sanders and Cunningham model (1972, 1973) as amplified by Lawson (1974).

While a second version of ROI was being created, a multiple-choice test covering the information in the game was written and validated. An attitude scale was selected to measure degree of player approval of large business organizations.

The two versions of ROI were then played by four pairs of classes in four different schools. In each school,

one class played the original version while the other class played the revised version. Controls were placed on teacher effect and information leakage among students.

Two research designs were employed. The first compared the pretest-to-posttest gain scores on the multiple choice test achieved by the "unevaluated game" players with those of the "revised game" players; the second was a post-test-only design to compare favorability in attitude toward large businesses by players of the two versions of ROI (Campbell & Stanley, 1963). The groups had been judged essentially similar on the basis of an analysis of demographic factors.

The results (Tables 1 and 2) were less than a ringing endorsement for the formative evaluation process employed. While differences were not significant, the students who played the original, "unevaluated" version of ROI actually achieved greater pretest-to-posttest cognitive gain scores and demonstrated a more favorable attitude toward business than the students who played the "revised" game.

Insert Table 1 about here

Insert Table 2 about here

The finding that an elaborate formative evaluation process failed to produce a significantly-improved version

Table 1

Comparison of Mean Gain Scores on ROI Multiple Choice Test

| Group | N | Pretest Mean | Posttest Mean | Mean Gain | SD | t | DF | Probability |
|---|---|---|---|---|---|---|---|---|
| "Revised" Game | 109 | 23.8257 | 27.0459 | 3.2202 | 5.054 | 0.32 | 212 | 0.746 |
| "Unevaluated" Game | 105 | 21.6571 | 25.1048 | 3.4476 | 5.187 | | | |

Table 2

Comparison of Mean Scores on "Favorability Toward Business" Questionnaire

| Group | N | Mean Score | SD | t | DF | Probability |
|---|---|---|---|---|---|---|
| "Revised" Game | 103 | 8.0078 | 1.923 | 0.56 | 200 | 0.573 |
| "Unevaluated" Game | 99 | 8.1626 | 1.973 | | | |

of Return on Investment may be attributed to several causes. The first, and most obvious, is the inadequacy of the particular formative evaluation model developed to revise the game. The lack of significant difference in learner performance produced by the two versions of ROI is eloquent evidence of this inadequacy. Conventional wisdom would direct the author back to the drawing board to resume the effort to create formative evaluation techniques which would produce significant differences and thus merit testing on a broader spectrum of instructional products.

It is possible to advance another explanation for the failure of the ROI formative evaluation model. This simply states that ROI was really a very "good" game to begin with and that subjecting it to a formative evaluation process represented little more than elaborate "tinkering around" which could not be justified by any analysis of costs versus benefits.

Before the creator of ROI plunges into an orgy of self-praise for his inordinate skill as a simulation game developer, however, prudence dictates that he consider this postulate: by the time any instructional simulation game reaches the point at which it can be said to be "playable" (substitute the word "useable" for other kinds of instructional products), further attempts to refine it by elaborate field testing are unlikely to be cost or time effective in

terms of improved learner performance.

Failure to design a model capable of validation is not surprising. The literature is replete with unvalidated formative evaluation models which are, by definition, ungeneralizable. Dick (1980) has noted the paucity of empirical studies of the effects of formative evaluation in instructional design. The appearance of a substantial body of such studies is long overdue after nearly a decade-and-a-half which has witnessed application of formative evaluation to instructional products. It is time to determine whether formative evaluation can be something more than a first draft strategy for instructional product development.

If formative evaluation models cannot be constructed which demonstrably contribute to improvement of learner performance, the utility of formative evaluation as an instructional development technique is questionable. If models are not verifiable, then formative evaluation is reduced to little more than a trial-and-error process to validate each individual product. Trial-and-error evaluation is not necessarily bad. It can have tangential benefits such as focusing attention on, and building support for, instructional product development (in much the same way that Presidential study commissions are designed to marshal public support for legislative initiatives) . . . but trial-and-error processes are shoddy materials with which to build a science of instruction.

REFERENCES

Abedor, A. J. Development and validation of a model explicating the formative evaluation process for multimedia self-instructional learning systems (Doctoral dissertation, Michigan State University, 1971). Dissertation Abstracts International, 1972, 32 4915A. (University Microfilms No. 72-8621)

Abedor, A. J. Second draft technology: development and field test of a model for formative evaluation of self-instructional multi-media learning systems. Viewpoints, 1972, 48(4), 9-44.

Baker, E. L. Formative evaluation of instruction. In W. J. Popham (Ed.), Evaluation in education: current applications. Berkeley, Calif.: McCutchan Publishing Corp., 1974.

Baker, E. L. & Alkin, M. C. Formative evaluation of instructional development: ERIC/AVCR annual review paper. AV Communication Review, 1973, 21(4), 389-418.

Bloom, B. S. Learning for mastery, 1968. In M. Mohan & R. E. Hull (Eds.), Individualized instruction and learning. Chicago: Nelson-Hall Co., 1974.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill, 1971.

Cambre, M. A. Historical overview of formative evaluation of instructional media products. Educational Communication and Technology Journal, 1981, 29(1), 3-25.

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally College Publishing Co., 1963.

Cunningham, D. J. Comments on the case studies of formative evaluation--the sources of information. Viewpoints, 1972, 48(4), 111-117.

Dick, W. Formative evaluation. In L. J. Briggs (Ed.), Instructional design: principles and applications. Englewood Cliffs, N.J.: Educational Technology Publications, 1977.

Dick, W. Formative evaluation in instructional development. Journal of Instructional Development, 1980, 3(3), 3-6.

Etlinger, L. E. Formative developmental evaluation: an in-house model for modular or individualized curricula (Doctoral dissertation, Southern Illinois University, 1975). Dissertation Abstracts International, 1976, 36, 7832A. (University Microfilms No. 76-13,236)

Gooler, D. D. Formative evaluation strategies for major instructional development projects. Journal of Instructional Development, 1980, 3(3), 7-11.

Lawson, T. E. Formative instructional product evaluation: instruments and strategies. Englewood Cliffs, N.J.: Educational Technology Publications, 1974.

Light, J. A., & Reynolds, L. J. Debugging product and testing errors: procedures for the formative evaluation of an individualized mathematics curriculum. Viewpoints, 1972, 48(4), 45-78.

Miree, L. F., & Kibler, R. J. The adaptation of formative evaluation procedures from mastery learning to non-mastery learning systems or "you really gotta wanna." Paper presented at the annual meeting of the American Educational Research Association, 1977. (ERIC Document Reproduction Service No. ED 150 152)

Nevo, D. A model for the utilization of formative evaluation in the process of developing instructional materials. Paper presented at the annual meeting of the American Educational Research Association, April 1976. (ERIC Document Reproduction Service No. ED 126 109)

Sanders, J. R. Comments on the case studies of formative evaluation--search for structure. Viewpoints, 1972 48(4), 119-132.

Sanders, J. R., & Cunningham, D. J. A structure for formative evaluation in product development. Bloomington, Indiana: EREL-Indiana University Press, 1972. (ERIC Document Reproduction Service No. ED 064 328)

Sanders, J. R., & Cunningham, D. J A structure for formative evaluation in product development. Review of Educational Research, 1973, 43(2), 217-236.

Schoen, F. E. The money game. American Youth Magazine, 1977, 16(1), 5-8, 28.

Schoen, F. E. The effectiveness of a formative evaluation model in the construction of a simulation game (Doctoral dissertation, Wayne State University, 1980). Dissertation Abstracts International, 1981, 41, 4272A. (University Microfilms No. 8107242)

Scriven, M. The methodology of evaluation. In R. E. Stake, AERA monograph series on curriculum evaluation no. 1. Chicago: Rand McNally and Co., 1967, 39-83.

Smith, N. L., & Murray, S. L. The status of research on models of product development and evaluation. Educational Technology, 1975, 15(3), 13-17.

Stake, R. E. Generalizability of program evaluation: the need for limits. Educational Product Report, 1969, 2(5), 39-41.

Stake, R. E. Responsive evaluation. 1972. (ERIC Document Reproduction Service No. ED 075 487)

Tennyson, R. D. The role of evaluation in instructional development. Educational Technology, 1976, 15(9), 17-24.

Thiagarajan, S. Learner verification and revision: what, who, when, and how? Audiovisual Instruction, 1976, 21(1), 18-19.